

Augmenting Image Classification Metrics with Tangents to Invariants: A Work in Progress

Investigators:

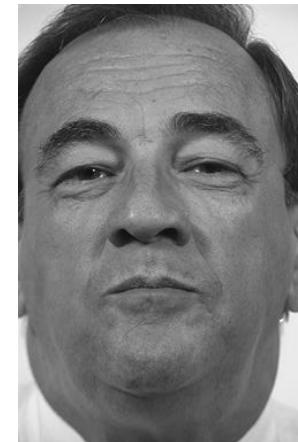
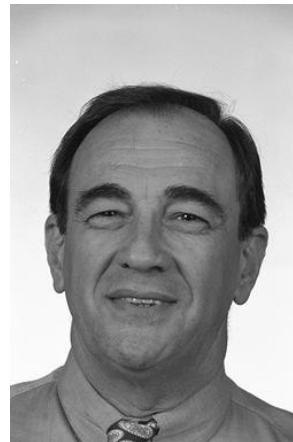
- Andy Fraser
- Nick Hengartner
- Brendt Wohlberg
- Kevin Vixie

Support:

- LANL LDRD
- PSU

Challenge: Accommodate Small *Shifts* of Images Along Invariant Manifolds Within a Maximum Likelihood Classification Framework

- Shift in x
 - Shift in y
 - Change in x scale
 - Change in y scale
 - Image rotation
-



- 3-d rotation
- Change in illumination



Prior Work and Inspiration

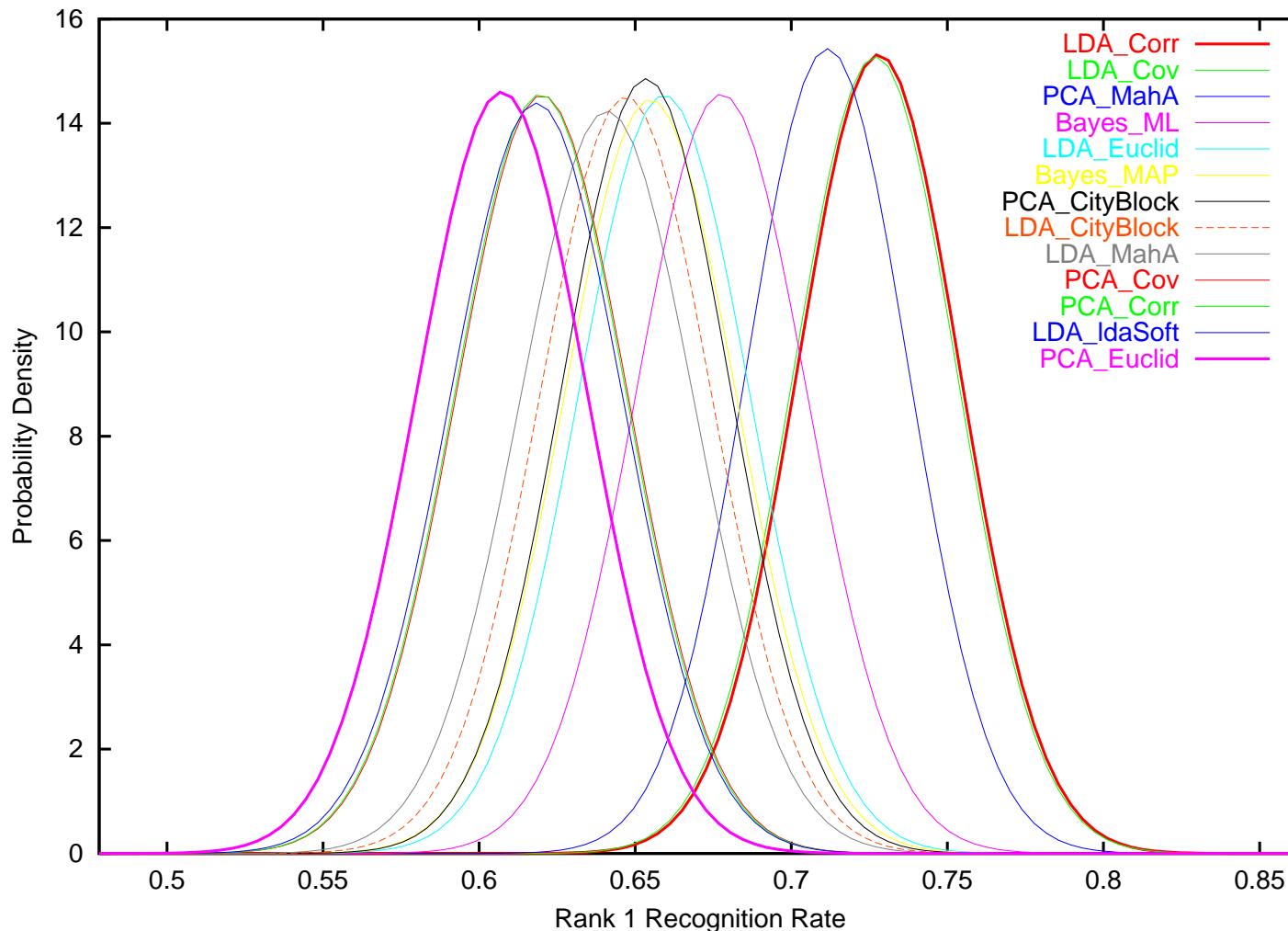
We collected and at least skimmed more than 160 papers. The following were especially helpful:

- Belhumeur, Hespanha, and Kriegman, “Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection.” 1996
- Simard, Le Cun, Denker, Victorri, “Transformation Invariance in Pattern Recognition — Tangent Distance and Tangent Propagation.” 1998
- Phillips, Wechsler, Huang, and Rauss, “The FERET Database and Evaluation Procedure for Face Recognition Algorithms.” 1998

CSU Test Bed

- Beveridge, Kai She, Draper, and Givens, “A nonparametric statistical comparison of principal component and linear discriminant subspaces for face recognition.” 2001
 - **Code available** at <http://www.cs.colostate.edu/evalfacerec/>
 - Designed to operate on **FERET** images (from NIST)
 - Image pre-processing code
 - CSU’s versions of PCA, PCA+LDA, and Bayesian Intrapersonal/Extrapersonal Difference classifier algorithms (many distance/angle measures; total of 13 classifiers)
 - Monte Carlo experiments on images with labeled eye locations that produces **distribution of recognition rates**

Algorithm Performance (CSU code)



Preprocessing

- Scale, shift, and rotate based on eye coordinates
- Elliptical mask
- Scale pixel values to equalize histogram
- Shift and scale pixel values to force Mean=0 & StdDev=1

Training

$591 = 197 \cdot 3$ training images

PCA Establish **projection basis B** of $354 = 0.6 \cdot (591 - 1)$ vectors

$$Y = [I_1, I_2, \dots, I_{591}]$$

$$X = Y - \bar{Y}$$

$$M = X^T \cdot X$$

$$V^T \Lambda V = M$$

$$b = X V^T$$

$$B = \text{normalized}(b)$$

LDA Establish **Fisher basis f** of $196 = (197 - 1)$ basis vectors

Project

$$Z = B \cdot X$$

From Z compute

$$\Sigma_w \text{ & } \Sigma_b$$

Find Fisher basis f

$$\text{Maximizes } \frac{|f \Sigma_b f^T|}{|f \Sigma_w f^T|}$$

LDA Calculate **within class covariance Σ_w**

PCA Euclidean Distance

A & B: Processed images projected onto PCA basis

Distance: $d(A, B) = \sqrt{\sum_k (A_k - B_k)^2}$

LDA Correlation

A & B : Processed images projected onto LDA basis

Means:

$$M_A = \frac{1}{n} \sum_k A_k, \quad M_B = \frac{1}{n} \sum_k B_k$$

Sums of squares:

$$V_A = \sum_k (A_k - M_A)^2, \quad V_B = \sum_k (B_k - M_B)^2$$

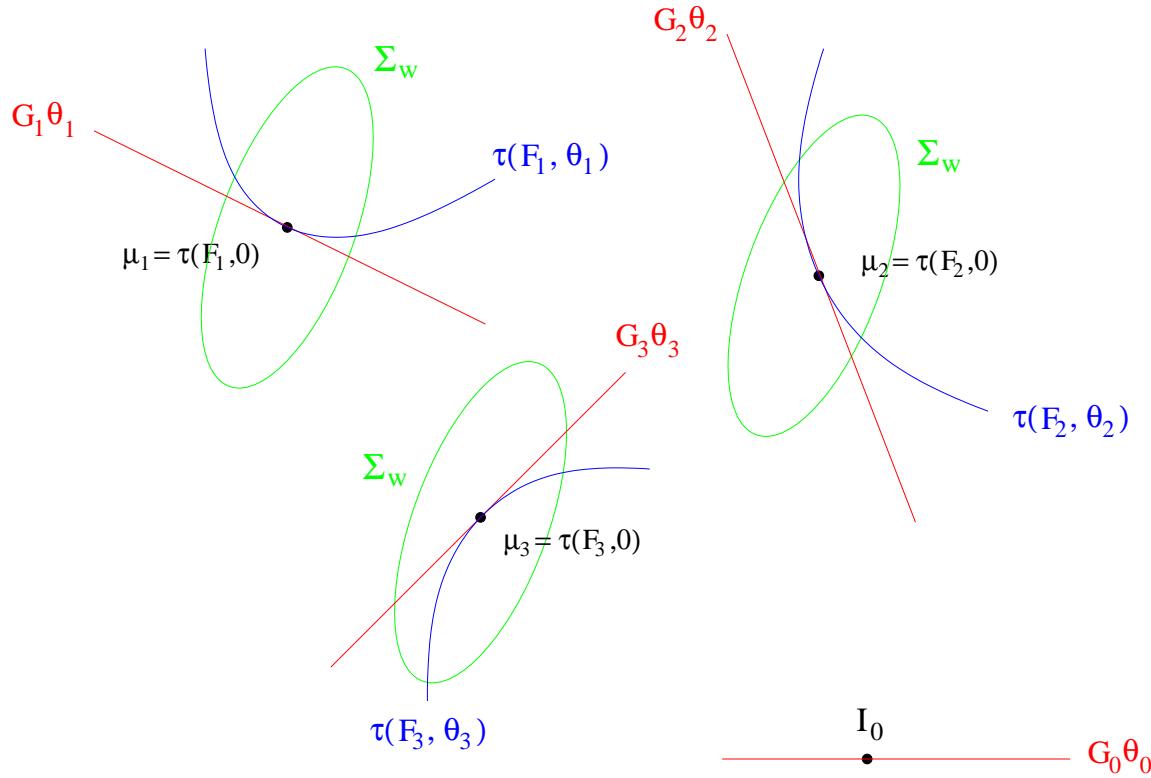
Cross terms:

$$C = \sum_k (A_k - M_A)(B_k - M_B)$$

Distance:

$$d(A, B) = 1 - \frac{C}{\sqrt{V_A V_B}}$$

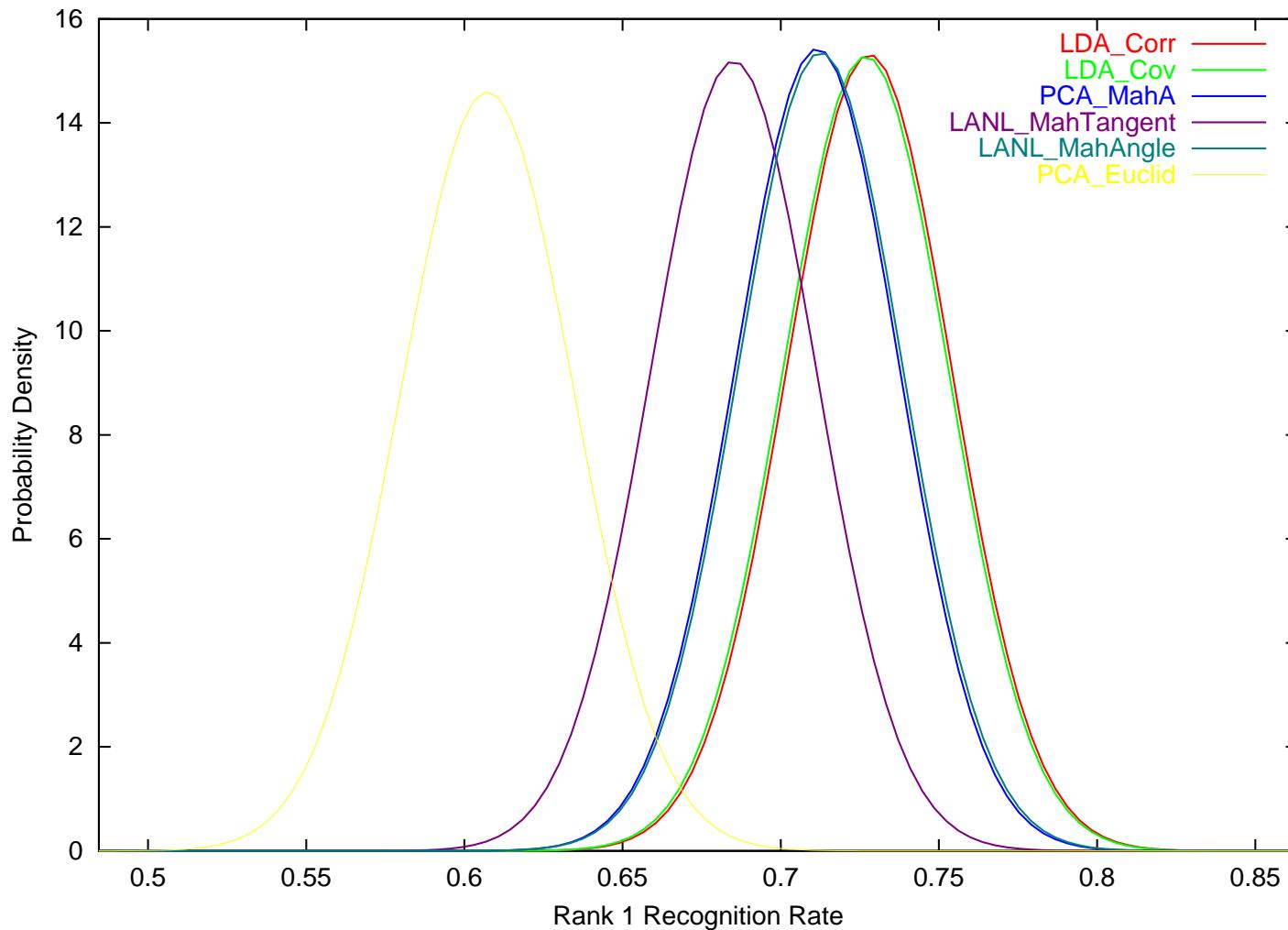
Tangents



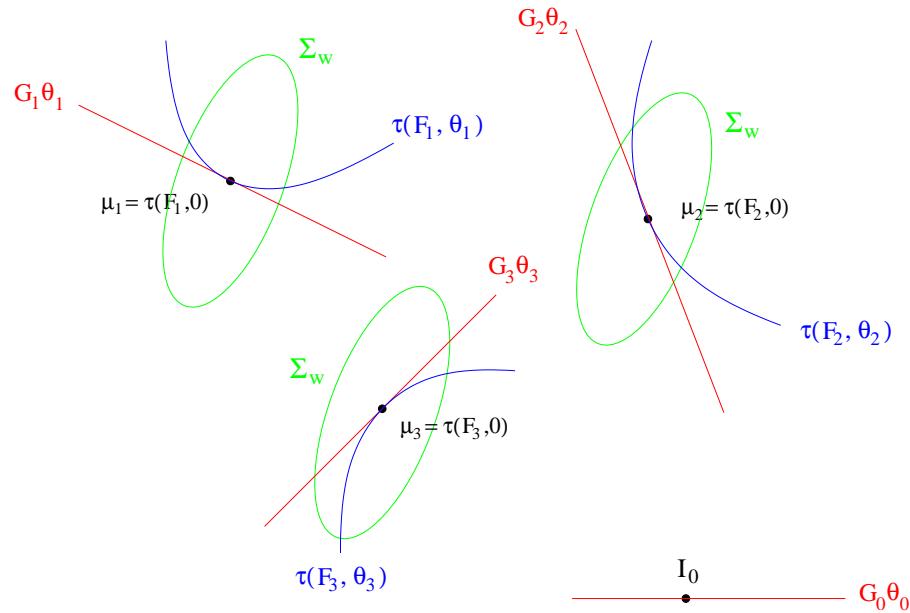
$$\hat{F}(I_0) = \operatorname{argmin}_i \left((I_0 - \mu_i)^T \Sigma_i^{-1} (I_0 - \mu_i) \right)$$

$$\Sigma_i = \Sigma_w + \alpha \mathbf{G} (\bar{H})^{-1} \mathbf{G}^T$$

Preliminary Results



Implementation Details



\mathcal{F}

Set of faces or individuals

Θ

Space of parameters for location on
an invariant manifold

I

Space of images

$$\tau: \mathcal{F} \times \Theta \mapsto I$$

Details cont: Normality

Ansatz:

$$I_i = \tau(F_i, \theta_i) + \varepsilon_i$$

$$\varepsilon_i \sim \mathcal{N}(0, \Sigma_w) \text{ and } \theta_i \sim \mathcal{N}(0, \Sigma_{\theta, F_i})$$

Taylor series:

$$\tau(F, \theta) = \tau(F, 0) + G\theta + \theta^T H \theta + R$$

where

$$G_i = \left. \frac{\partial \tau(F, \theta)}{\partial \theta_i} \right|_{\theta=0} \text{ and } H_{i,j} = \left. \frac{\partial^2 \tau(F, \theta)}{\partial \theta_i \partial \theta_j} \right|_{\theta=0}$$

If $\theta_F \sim \mathcal{N}(0, \Sigma_{\theta_F})$:

$$G\theta_F \sim \mathcal{N}(0, G\Sigma_{\theta_F} G^T),$$

Details cont: First Order Approx

$\theta^T H \theta + R \text{ small} \implies I_F$ approximately normal:

$$\begin{aligned} I_F &\sim \mathcal{N}(\tau(F, 0), \textcolor{red}{G}\Sigma_{\theta_F} G^T + \textcolor{green}{\Sigma_w}) \\ &\equiv \mathcal{N}(\mu_F, \Sigma_F) \end{aligned}$$

Maximum likelihood classification:

$$\begin{aligned} \hat{F}(I) &= \operatorname{argmin}_i (I - \mu_i)^T \textcolor{magenta}{\Sigma_i^{-1}} (I - \mu_i) \\ \Sigma_i &= \textcolor{green}{\Sigma_w} + \textcolor{red}{G_i}\Sigma_{\theta} G_i^T \end{aligned}$$

Details cont: Large and Small

Conflicting goals:

1. Want Σ_θ large. Classify images with large invariant shifts
2. Want $\theta^T H \theta \in I$ small. So Taylor series truncation error small

Defining:

$$\sum_{p=1}^N \theta^T \sqrt{H_p H_p} \theta \equiv \theta^T H_{\text{mrs}} \theta$$

Maximize the determinant: $|\Sigma_\theta|$

Subject to: $\mathbb{E} \theta^T H_{\text{mrs}} \theta \leq K$

Solution:

$$\Sigma_\theta = \alpha H_{\text{mrs}}^{-1}$$

Value of α balances competing goals

Details cont: Improvement from Diagonalizing Σ_w

In terms of the decomposition

$$\Sigma_w = \sum_d \mathbf{e}_d \lambda_d \mathbf{e}_d^T$$

and the $k \times N \times k$ tensor

$$\mathbf{H}$$

define the $k \times k$ matrix

$$\mathbf{H}_d \equiv (\mathbf{H} \mathbf{e}_d)$$

and the $k \times k$ matrix

$$\bar{\mathbf{H}} \equiv \sum_d \lambda_d^{-\frac{1}{2}} \sqrt{(\mathbf{H}_d)^T \mathbf{H}_d}.$$

Setting

$$\Sigma_\theta = \alpha \bar{\mathbf{H}}^{-1}$$

allows displacements that produce large truncation errors along directions that images of individuals typically vary.

Derivatives

$$\tau(F, 0)(x, y) = \int \textcolor{red}{I}(u, v) \textcolor{blue}{K}(x - u, y - v) du dv$$

eg, $\theta_1 = x$ shift

$$\begin{aligned} \tau(F, \theta)(x, y) &= \int \textcolor{red}{I}(u + \theta_1, v) \textcolor{blue}{K}(x - u, y - v) du dv \\ &= \int \textcolor{red}{I}(t, v) \textcolor{blue}{K}(x - t - \theta_1, y - v) dt dv \end{aligned}$$

$$\frac{\partial \tau(F, \theta)}{\partial \theta_1} \Big|_{\theta=0} (x, y) = \int \textcolor{red}{I}(u, v) \left(\frac{\partial K}{\partial x} \right)_{(x-u, y-v)} du dv$$

Kernel K is Gaussian, and we use FFT for convolution

Derivatives, cont.

For more general θ , write

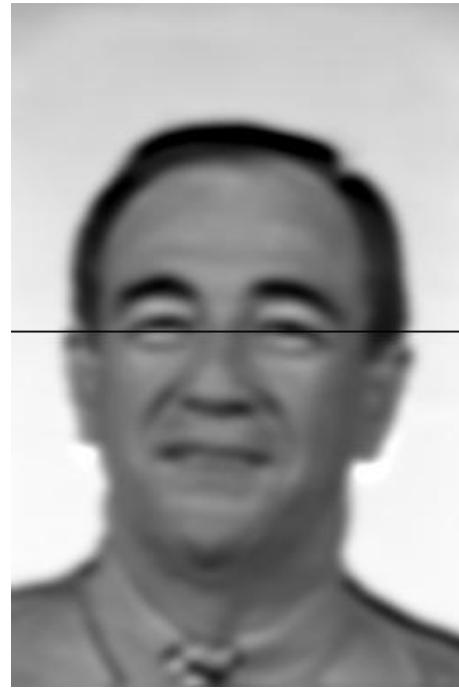
$$\begin{aligned}\tau(F, \theta)(z) &= \int \textcolor{red}{I}(T_\theta(w)) \textcolor{blue}{K}(w - z) dw \\ &= \int \textcolor{red}{I}(u) \textcolor{blue}{K}(T_\theta^{-1}(u) - z) \left[(D_w T_\theta)_{T_\theta^{-1}(u)} \right]^{-1} du\end{aligned}$$

As with θ_1 , first and second derivatives are convolutions that we evaluate with FFTs.

Shifts



Image **convolved** with
kernel



First order
approximation to **shift**



Second order
approximation to **shift**